Quantum Anomalous Hall Effect of Magnetic Topological Insulator Thin Films by Phase Boundary Engineering

Kiminori Hattori

Graduate School of Engineering Science, Osaka University, Toyonaka, Osaka 560-8531, Japan

We generalize the topological phase diagram of magnetic topological insulator (TI) thin films in an extended parameter space comprising out-of-plane (OP) and in-plane (IP) exchange fields in the presence of structural inversion asymmetry (SIA) while taking a generic orbital-dependent spin coupling allowed for TIs into consideration. The results show that an IP field substantially deforms phase boundaries and generically induces the quantum anomalous Hall (QAH) effect. For symmetric spin coupling, extremely weak OP and IP exchange fields create the QAH state by tuning SIA with a gate bias. For antisymmetric coupling, the QAH phase is absent without a strong enough IP field. These findings demonstrate that in the thin-film regime, engineering the phase boundary is a key process to efficiently realize and manipulate the QAH effect for nondissipative electronic applications.

I. INTRODUCTION

A thin-film topological insulator (TI) provides a planar platform that supports a variety of topological states and is suitable for exploiting related topological effects in actual devices. In a thin-film configuration, surfaces states of a three-dimensional (3D) TI[1–4] lying on the opposite surfaces hybridize to create a tunneling gap. Then, the system alternates between the topologically-trivial normal insulator (NI) phase and the nontrivial quantum spin Hall (QSH) phase, depending on the film thickness.[5–12] The QSH state consists of two copies of quantum Hall (QH) states with opposite Chern numbers, arranged so that time-reversal symmetry is preserved.[13–16] In a finite system with open boundaries, there exists on each edge a Kramers pair of chiral edge states inside the gap. The helical edge mode is essentially robust against time-reversal-invariant nonmagnetic perturbations. However, magnetic disorder due to random magnetic impurities is detrimental to the QSH transport since they give rise to intraedge backscattering.[16]

This vulnerability can be eliminated by incorporating ferromagnetism into the system by magnetic doping.[17–21] A strong enough exchange field normal to the surface drives the NI or QSH phase into the quantum anomalous Hall (QAH) phase, where the QH effect is activated in either of two subsystems and suppressed in the other.[22–26] The QAH phase hosts a single chiral edge mode on each edge, for which backscattering is completely forbidden. Hence, the QAH insulator exhibits perfect conduction that is insensitive to magnetic and nonmagnetic disorder as long as the gap remains intact.[16] The QAH effect is experimentally observed in the Bi_2Se_3 family of materials[27–30], and is theoretically proposed to occur in other related systems[31–35].

There are extensive studies in this context, where an out-of-plane (OP) and orbital-independent exchange field constitutes a main control parameter to induce the QAH effect for a given hybridization gap. Here, we generalize the topological phase diagram of magnetic TI thin films by considering an in-plane (IP) exchange field in addition to the OP field, structural inversion asymmetry (SIA) due to an OP electric field, and generic orbital-dependent exchange coupling for the Dirac model describing 3DTIs. It is shown in the extended parameter space that an IP exchange field generically induces the QAH effect by appreciably deforming phase boundaries, leading to an unconventional scheme to control the topological state of a thin-film TI.

II. THEORY

The 4×4 Dirac Hamiltonian describing nonmagnetic 3DTIs is given in momentum space by

$$\mathcal{H}_0 = D(\mathbf{k}) + \mathbf{A}(\mathbf{k}) \cdot \tau_x \boldsymbol{\sigma} + B(\mathbf{k})\tau_z, \qquad (1)$$

where $D = D_1 k_z{}^2 + D_2 k^2$, $\mathbf{A} = (A_2 k_x, A_2 k_y, A_1 k_z)$, $B = B_0 - B_1 k_z{}^2 - B_2 k^2$, and $k^2 = k_x{}^2 + k_y{}^2$.[1,2] The parameters in these definitions are estimated for the Bi_2Se_3 family.[1,2] σ_μ and τ_μ with $\mu = x, y, z$ represent the Pauli matrices in spin and orbital spaces, respectively. To formulate the exchange interaction with magnetic moments in a homogeneous phase, we consider the generic form of Zeeman-type coupling allowed in the Dirac Hamiltonian[2,20]

$$\mathcal{H}_m = (\mathbf{m} + \mathbf{n}\tau_z) \cdot \boldsymbol{\sigma}. \qquad (2)$$

The exchange field (in energy units) is composed of the symmetric part $\mathbf{m}$ and the antisymmetric part $\mathbf{n}$, by which the possible orbital dependence of spin coupling is fully describable. Thus, the total Hamiltonian consists of $\mathcal{H} = \mathcal{H}_0 + \mathcal{H}_m$. See Appendix A for more details. We first analyze the contribution due to the symmetric field $\mathbf{m}$, which accords with the conventional approach used in the study of magnetic TIs. The effect of the antisymmetric counterpart $\mathbf{n}$ will be addressed later.

Diagonalizing $\mathcal{H}_0$ under an open boundary condition and projecting $\mathcal{H}$ onto the subspace of surface states,[1,2,7,8,11,12,23–25] we obtain the effective low-energy Hamiltonian in a thin-film configuration. The result is

summarized as $H = H_0 + H_m$ with

$$H_0 = \varepsilon(\mathbf{k}) + \gamma s_z(\boldsymbol{\sigma} \times \mathbf{k})_z + t(\mathbf{k})s_x + m_z\sigma_z, \quad (3)$$

$$H_m = (I + Js_x)\mathbf{m}_{||} \cdot \boldsymbol{\sigma}, \quad (4)$$

in the basis constituted of top and bottom surface states $\{t_\uparrow, t_\downarrow, b_\uparrow, b_\downarrow\}$. Here, s_μ denotes the Pauli matrix in orbital space spanned by $\{t, b\}$. H_m describes the exchange interaction associated with the IP field $\mathbf{m}_{||} = (m_x, m_y, 0)$. The contribution due to the OP field m_z is incorporated into H_0 for convenience. H_0 preserves rotational symmetry while it is broken by H_m. The parameters are given by $\varepsilon = E_0 + Dk^2$, $t = \Delta - Bk^2$, $D = D_2 - IB_2$, $B = JB_2$, $I = \frac{\langle t|\tau_z|t\rangle + \langle b|\tau_z|b\rangle}{2}$, and $J = \mathrm{Re}\,\langle t|\tau_z|b\rangle$. By definition, $|J|$ reflects intersurface mixing, and is essentially augmented with a reduction in the film thickness L_z. For instance, $(I, J) = (0.062, -0.7)$ at $L_z = 2.3$nm and $(0.12, -0.35)$ at 3nm for Bi$_2$Se$_3$,[12] whereas $I = D_1/B_1 = 0.13$ and $J = 0$ in the thick-film limit $L_z \to \infty$.[1,2,8,25] This suggests that in the thin-film regime, the J term is dominant in H_m.

The orthogonal transformation $|\varphi_{1,2}\rangle = (|t\rangle \pm |b\rangle)/\sqrt{2}$ arranges H_0 into a block diagonal form. The eigenenergies of two blocks labeled by $p = \pm 1$ are written as $E_r^{(p)} = \varepsilon + rE^{(p)}$ and $E^{(p)} = \sqrt{\gamma^2 k^2 + (t + pm_z)^2}$, where $r = \pm 1$ refer to the conduction and valence bands, respectively. The band gap $E_g = 2\min|\Delta \pm m_z|$ closes at the Γ point ($k = 0$) when $m_z^2 = \Delta^2$. On the other hand, it is shown from the Chern numbers of the associated four bands that block p becomes nontrivial when $(\Delta + pm_z)B > 0$.[7,8,11,12,23,24] As a result, the system enters into the QAH phase for $m_z^2 > \Delta^2$. If this criterion is not fulfilled, the system lies in the QSH phase for $\Delta B > 0$ or the NI phase for $\Delta B < 0$. These three phases are topologically distinct in terms of the Chern numbers and cannot smoothly connect without closing the gap.

The total Hamiltonian $H = H_0 + H_m$ can be also analytically diagonalized for $I = 0$. Defining the orthogonal two wavevectors $\mathbf{k}_\ell = \mathbf{m}_{||}(\mathbf{m}_{||} \cdot \mathbf{k})/m_{||}^2$ and $\mathbf{k}_t = \mathbf{k} - \mathbf{k}_\ell$, the gap function is expressed as

$$E^{(p)} = \sqrt{(T + pM)^2 + \gamma^2(k^2 - \eta k_\ell^2)}, \quad (5)$$

where $T = \sqrt{t^2 + \eta\gamma^2 k_\ell^2}$, $M = \sqrt{m_z^2 + J^2 m_{||}^2}$, and $\eta = \frac{J^2 m_{||}^2}{m_z^2 + J^2 m_{||}^2}$. It is found from Eq. (5) that the gap closes when $T - M = (1 - \eta)k_\ell = k_t = 0$. These coupled conditions reduce to $k = 0$ and $M^2 = \Delta^2$ for $m_z \neq 0$, and predict that the QAH phase emerges in the region where $M^2 > \Delta^2$. The system otherwise belongs to the QSH phase for $\Delta B > 0$ or the NI phase for $\Delta B < 0$. Thus, adding H_m to H_0 is formally equivalent to replacing m_z^2 by M^2 in terms of gap closing, demonstrating that m_z and $m_{||}$ equally drive the phase transition into the QAH state. If $m_z = 0$, the gap remains closed as long as $J^2 m_{||}^2 > \Delta^2$, and then the system becomes metallic. From these considerations, we expect the Chern number

to be formulated as

$$C = -\mathrm{sgn}(m_z)\theta(m_z^2 + J^2 m_{||}^2 - \Delta^2). \quad (6)$$

It is worth noting that two QAH phases classified by $C = \pm 1$ touch for $J^2 m_{||}^2 > \Delta^2$ at $m_z = 0$ where the gap vanishes. In the QSH phase, the bulk gap is open even for $m_{||} \neq 0$, although helical edge states formed in a finite system are gapped by intraedge mixing. In this sense, the system under a small IP field is no longer in the ideal QSH state. In the following, we nevertheless refer to it as a QSH system to avoid unnecessary and probably confusing nomenclature. For more details, see Appendix B.

A similar phase transition arising from an IP field has been investigated previously for paramagnetic HgMnTe quantum wells,[36] although its physical mechanism remains unresolved. To understand the underlying physics, we here analyze the effective Hamiltonian for a weak IP field. The relevant Hamiltonian is formulated as $\tilde{H} = H_0 + H_m(E_r^{(p)} - H_0)^{-1}H_m$ for an unperturbed eigenstate of energy $E_r^{(p)}$ following the second-order perturbation formalism. $\tilde{H}$ is block diagonal in the basis chosen for H_0. An identity term in each block is negligible since it has no effect on the band topology. The leading terms in two blocks construct the rotationally-invariant effective Hamiltonian, expressed as

$$\tilde{H} = \gamma s_z(\boldsymbol{\sigma} \times \mathbf{k})_z + \tilde{t}(\mathbf{k})s_x + \tilde{m}_z\sigma_z, \quad (7)$$

yielding the renormalized parameters to second order in k; $\tilde{t} = \tilde{\Delta} - \tilde{B}k^2$, $\tilde{\Delta} = \Delta - J^2 m_{||}^2/4\Delta$, $\tilde{B} = B(1 + J^2 m_{||}^2/4\Delta^2)$, and $\tilde{m}_z = m_z + J^2 m_{||}^2/4m_z$. It is easily shown from these formulae that the original QAH criterion $m_z^2 > \Delta^2$ substantially changes into $m_z^2 + J^2 m_{||}^2 > \Delta^2$ to second order in $m_{||}$. The renormalization thus reproduces the conclusion drawn from gap closing.

It is also instructive to consider a generalized 3D parameter space $(\Delta, m_z, m_{||})$. Scanning a 2D section (Δ, m_z) along $m_{||}$, the renormalization due to $m_{||}$ is viewed as a continuous deformation of the QAH phase boundary, which consists of an equilateral hyperbola and a line segment joining two vertices at $(\pm J m_{||}, 0)$. As $m_{||}$ increases, the QAH effect is activated in the otherwise inactive region by deforming the phase boundary outward, i.e., $m_{||}$ intrinsically extends the QAH phase. However, the renormalization is no longer valid in the thick-film limit where $\Delta = J = 0$. In this limit, top and bottom surface states are completely decoupled, and $\mathbf{m}_{||}$ merely shifts each Dirac dispersion in the opposite direction by an amount $s_z I(\mathbf{m}_{||} \times \hat{\mathbf{z}})/\gamma$. This is a clear distinction showing that the renormalization and the consequent phase-boundary deformation are characteristic of thin-film TIs.
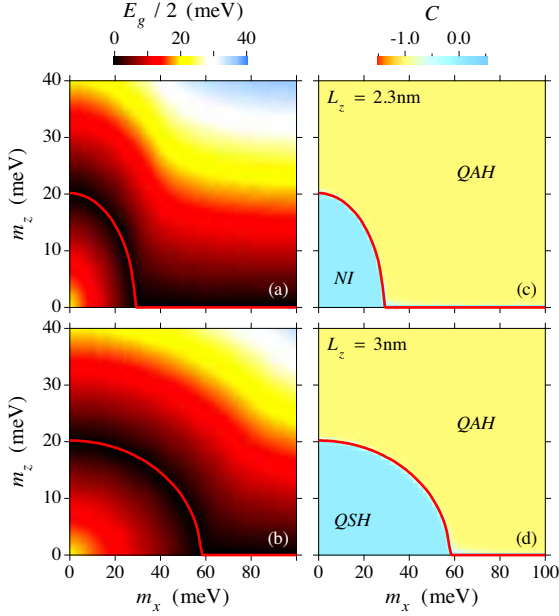
FIG. 1. (Color online) [(a) and (b)] Band gap E_g and [(c) and (d)] Chern number C numerically computed as a function of symmetric exchange fields m_x and m_z. In the calculation, SIA is neglected ($V = 0$). Solid lines represent the phase boundaries by which the QAH phase is separated from the NI phase for [(a) and (c)] $L_z = 2.3$nm and from the QSH phase for [(b) and (d)] $L_z = 3$nm.

III.   NUMERICAL CALCULATION

To examine quantitative aspects, we have performed numerical calculation of the Chern number using the Kubo formula $C = \frac{1}{2\pi}\sum_{n\in\text{occ}}\int_{\text{BZ}} d\mathbf{k}\,\Omega_n$ for a periodic system,[37] where the summation runs over all occupied states at half filling, and the integration is over the Brillouin zone. The momentum-space Berry curvature is given by

$$\Omega_n(\mathbf{k}) = 2\,\text{Im}\sum_{m\neq n}\frac{\langle n|\frac{\partial H}{\partial k_x}|m\rangle\langle m|\frac{\partial H}{\partial k_y}|n\rangle}{(E_n - E_m)^2}, \qquad (8)$$

where $H|n\rangle = E_n|n\rangle$. For simplicity, we set $\mathbf{m}_{||} = (m_x, 0, 0)$ without loss of generality because H_0 is rotationally invariant. The film thicknesses were chosen as $L_z = 2.3$nm and 3nm, for which equally $|\Delta| = 20.2$meV. In the absence of exchange coupling, the 2.3nm- and 3nm-thick systems are in the NI phase ($\Delta B < 0$) and in the QSH phase ($\Delta B > 0$), respectively. Further details are described in our previous paper.[12] In the calculation, the numerical values of I and J evaluated from surface wavefunctions were adopted.

The numerical results are shown in Fig. 1. As expected, the gap-closing line along $m_z{}^2 + J^2 m_{||}{}^2 = \Delta^2$ for $J^2 m_{||}{}^2 < \Delta^2$ and $m_z = 0$ for $J^2 m_{||}{}^2 > \Delta^2$ separates the QAH phase featured by $C = -\text{sgn}(m_z)$ from the NI or QSH phases carrying $C = 0$. In the region where

$m_z{}^2 < \Delta^2$, a strong enough $m_{||}$ brings about the topological phase transition into the QAH phase. The formation of the QAH state is independently verified from chiral edge states in an open-boundary system and the resulting quantization of Hall resistance (not shown). The system stays in the QAH phase for $m_z{}^2 > \Delta^2$ irrespective of m_x. In this regime, the band gap continuously increases with increasing m_x. The QAH state is topologically protected and essentially immune to external perturbations unless the gap collapses. In this sense, the QAH effect is stabilized by adding an IP field through the gap augmentation.

Next, we will take SIA stemming from an OP electric field into consideration. The relevant Hamiltonian is simply given by[8,11,24]

$$H_v = V s_z. \qquad (9)$$

In the absence of H_m, the total Hamiltonian $H = H_0 + H_v$ is easily diagonalized. The gap function is given by $E^{(p)} = \sqrt{(T + pM)^2 + (1 - \eta)\gamma^2 k^2}$, where $T = \sqrt{t^2 + V^2}$, $M = \sqrt{m_z{}^2 + \eta\gamma^2 k^2}$, and $\eta = \frac{V^2}{t^2 + V^2}$. It is found from the gap-closing condition $T - M = (1 - \eta)k = 0$ that the QAH phase emerges when $m_z{}^2 > \Delta^2 + V^2$, indicating that SIA tends to suppress the QAH effect.[24] This can be explicitly shown in (Δ, m_z) space from the phase-boundary deformation due to V. The QSH phase exists for $\Delta B > 0$ within the interval $V^2 - V_0{}^2 < m_z{}^2 < \Delta^2 + V^2$. Here, $V_0 = |\gamma|k_0$, and $k_0 = \sqrt{\Delta/B}$ is the critical momentum at which $\eta = 1$.[11,24] The NI phase appears otherwise, i.e., $m_z{}^2 < V^2 - V_0{}^2$ for $\Delta B > 0$ and $m_z{}^2 < \Delta^2 + V^2$ for $\Delta B < 0$.

Figure 2 summarizes the numerical results for $L_z = 2.3$nm in the (V, m_z) plane for various m_x values. Solid lines in the figure indicate the points where the gap vanishes. As is clearly seen, the QAH phase gradually extends with increasing m_x, and eventually covers the entire parameter space. A similar behavior is also observed for $L_z = 3$nm in Fig. 3. It is confirmed from these observations that even in the presence of SIA, an IP field largely deforms phase boundaries and induces the QAH effect independently of the initial topology.

A remarkable feature captured in Fig. 3 is penetration of the QAH phase through the original QSH-NI boundary in the presence of a small m_x. To elucidate this peculiar behavior more quantitatively, the phase diagram in (V, m_x) space at $m_z = 10$meV ($< |\Delta|$) is displayed in Fig. 4. In this diagram, the phase boundary $m_x(V)$ amounts to the critical field to generate the QAH effect. For $L_z = 2.3$nm, $m_x(V)$ monotonically increases as a function of V from the finite lower bound given by $m_x(0) = \sqrt{(\Delta^2 - m_z{}^2)/J^2}$. On the other hand, $m_x(V)$ decreases along the QSH-QAH boundary and subsequently increases along the NI-QAH boundary for $L_z = 3$nm. Importantly, $m_x(V)$ vanishes at the intermediate critical point where these three phases contact. In terms of the analytical results described above, such a triple point appears for $\Delta B > 0$ at $m_{||} = 0$ and
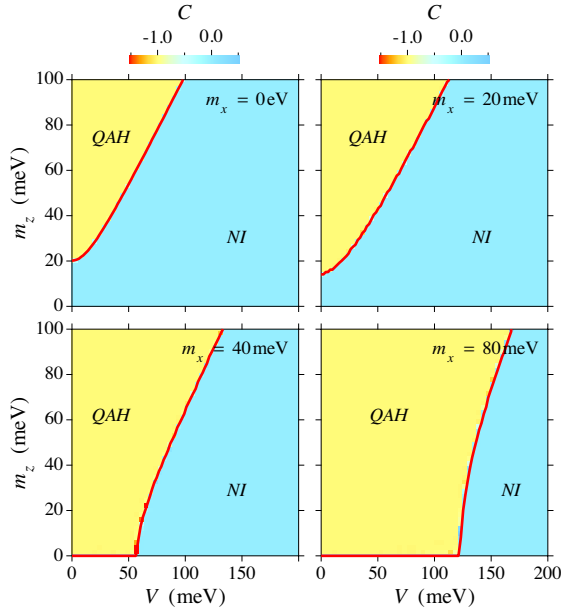
FIG. 2. (Color online) Numerical values of Chern number for $L_z = 2.3$nm as a function of V and m_z at (a) $m_x = 0$, (b) 20meV, (c) 40meV, and (d) 80meV. Solid lines represent the phase boundaries where the gap closes.

FIG. 3. (Color online) Numerical values of Chern number for $L_z = 3$nm as a function of V and m_z at (a) $m_x = 0$, (b) 20meV, (c) 40meV, and (d) 60meV. Solid lines represent the phase boundaries where the gap closes.

$V^2 = V_0{}^2 + m_z{}^2$, indicating that the QAH effect is inducible by extremely small m_z and $m_{||}$ in a thin-film geometry. This property presents a striking contrast to the prerequisite $m_z{}^2 > \Delta^2 + V^2$ for $m_{||} = 0$,[24] which implies that a larger gap $|\Delta|$ due to a stronger intersurface mixing obstructs the phase transition into the QAH state. The difficulty incidental to thickness reduction is circumvented by introducing a weak IP field. It is also noticeable that around the triple point, the topological state of a thin-film TI is very sensitive to V and is therefore tunable by a moderate gate bias. The tunablity is intrinsic to the thin-film regime and is not easy to realize in thick films.

IV. ANTISYMMETRIC EXCHANGE COUPLING

Finally, we examine the effects of antisymmetric exchange interaction. Spin coupling of this type adds the term

$$H_n = (I + Js_x)n_z\sigma_z + \mathbf{n}_{||} \cdot \boldsymbol{\sigma}, \qquad (10)$$

to the thin-film Hamiltonian. The antisymmetric fields In_z and $\mathbf{n}_{||}$ can be absorbed into definitions of the symmetric counterparts m_z and $I\mathbf{m}_{||}$, respectively. The remaining term related to Jn_z is a new contribution. Similar to $m_z + In_z$, Jn_z breaks time-reversal and reflection symmetries. The symmetry breaking is necessary for the QAH effect.[38] Nevertheless, Jn_z primarily impedes the QAH effect. This can be explicitly shown by assuming
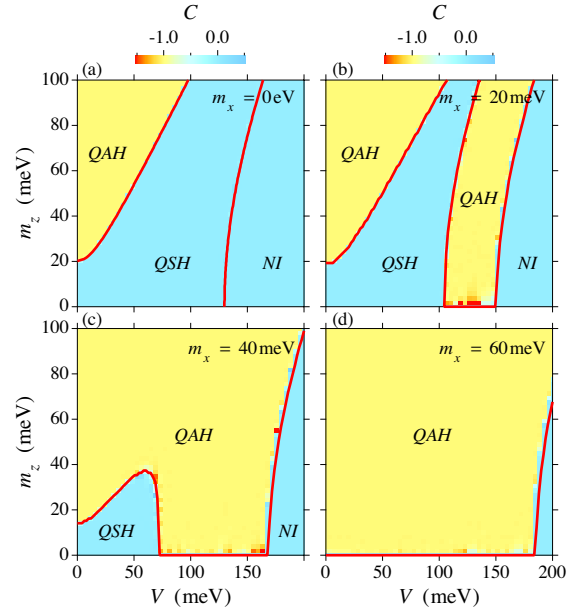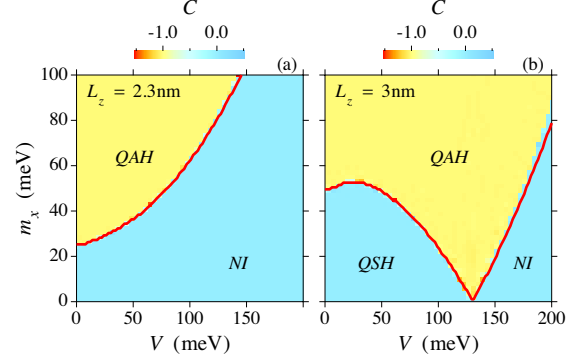

FIG. 4. (Color online) Numerical values of Chern number at $m_z = 10$meV as a function of V and m_x for (a) $L_z = 2.3$nm and (b) $L_z = 3$nm. Solid lines represent the phase boundaries where the gap closes.

$\mathbf{m}_{||} = \mathbf{n}_{||} = 0$. In this simplified situation, Jn_z gives rise to only an energy splitting between two blocks, $E_0{}^{(p)} = E_0 + pJn_z$, and consequently the system becomes metallic when Jn_z is sufficiently large.[20] The system remains in an insulating state for $|Jn_z| < \max(|\Delta|, |m_z + In_z|)$, whereas band inversion (gap closing and reopening) occurs in either of the two blocks when $|\Delta| < |m_z + In_z|$. Combining these two conditions, we reach the effective QAH criterion $|Jn_z| < |m_z + In_z|$, which conflicts with the relation $|J| > I$ intrinsic to thin films if $m_z = 0$. In thick films, $I > |J| \approx 0$ so that the surface QAH effect is realizable even for $m_z = 0$. However, edge states are no longer robust to external perturbations because of
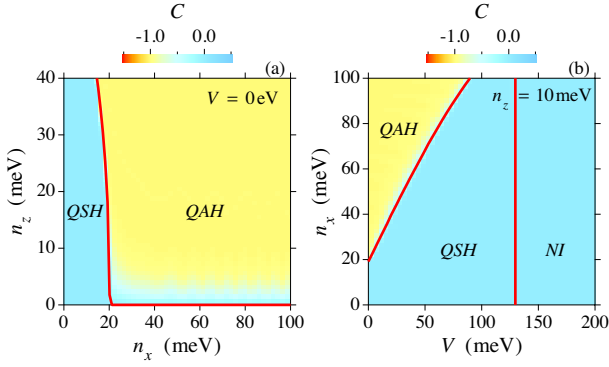
FIG. 5. (Color online) Numerical values of Chern number C for $L_z = 3$nm in the case of antisymmetric exchange coupling: (a) $C(n_x, n_z)$ at $V = 0$ and (b) $C(V, n_x)$ at $n_z = 10$meV. Solid lines represent the phase boundaries where the gap closes.

mixing with nonchiral side-surface modes.[29]

The disappearance of the QAH effect in the thin-film regime is reversed by introducing an IP field $n_{||}$, as shown in Fig. 5. Fundamentally, this effect is accountable for by the block mixing due to $n_{||}$, which substantially changes the QAH criterion into $J^2 n_z{}^2 + n_{||}{}^2 > \Delta^2$ for $\mathbf{m} = 0$. As demonstrated above, SIA tends to quench the QAH effect for symmetric coupling. This is also valid for antisymmetric coupling. In this case, it is shown from the analytical diagonalization for $V \neq 0$ and $n_{||} = 0$ that the QAH phase is absent; the system lies in either the QSH phase for $V_0{}^2 > V^2 + J^2 n_z{}^2$ and $\Delta B > 0$ or the NI phase otherwise. However, a sufficiently large $n_{||}$ regenerates the QAH effect in the presence of SIA, although the triple point is not located around $n_{||} = 0$ for a small n_z. Thus, a nonzero IP field is essential to realize the QAH effect, particularly for antisymmetric coupling. This observation is important in realistic situations since $|n_z|$ is estimated to be comparable to, or larger than, $|m_z|$ in a certain class of magnetic TIs.[20]

V. SUMMARY

In summary, we have generalized the topological phase diagram of magnetic TI thin films by considering IP and OP exchange fields, SIA induced by an OP electric field, and symmetric and antisymmetric spin couplings with magnetization, on the basis of the Chern number computed in a multidimensional parameter space. The results demonstrate that the phase-boundary deformation due to an IP exchange field generically induces the QAH effect. For symmetric coupling, the QAH state is created by minimal OP and IP exchange fields in the proximity of the QSH-QAH-NI triple point. For antisymmetric coupling, the QAH phase is absent without a strong enough IP field.

Even when magnetocrystalline anisotropy is perpendicular to the film, as observed for a large class of mag-

netic TI materials,[17,19–21] an IP exchange field can be generated by applying an external magnetic field $B_{||}$ parallel to the film, which rotates the magnetization by an angle $\theta = \sin^{-1} B_{||}/B_K$, where B_K denotes the anisotropy field. The auxiliary external field is suitable to separate exchange or Zeeman coupling from orbital effects in experiments. Moreover, in terms of the anisotropy energy to second order $E = K_1 \sin^2 \theta + K_2 \sin^4 \theta$, the magnetization canted by an angle $\theta = \sin^{-1} \sqrt{-K_1/2K_2}$ to the easy axis may form a ground state and stabilize with no external field by virtue of shape or surface anisotropy in a thin-film geometry.[39,40]

Appendix A: Bulk Hamiltonian

Here, we compare in more detail the bulk Hamiltonians, $\mathcal{H}_0$ and $\mathcal{H}_m$, to those presented in Refs. 1 and 2. $\mathcal{H}_0$ is identical to Eq. (1) in Ref. 1, as well as to Eq. (19) in Ref. 2 to second order in k. They are equally represented as

$$\mathcal{H}_0 = D(\mathbf{k}) + A_1 \Gamma_3 k_z + A_2 (\Gamma_1 k_x + \Gamma_2 k_y) + B(\mathbf{k})\Gamma_5,$$

where the Γ matrices are defined as $\Gamma_1 = \tau_x \sigma_x$, $\Gamma_2 = \tau_x \sigma_y$, $\Gamma_3 = \tau_x \sigma_z$, $\Gamma_4 = \tau_y$ and $\Gamma_5 = \tau_z$. $\mathcal{H}_0$ is unitary equivalent to Eq. (16) in Ref. 2, expressed as

$$\mathcal{H}'_0 = D(\mathbf{k}) + A_1 \Gamma_4 k_z + A_2 (\Gamma_1 k_y - \Gamma_2 k_x) + B(\mathbf{k})\Gamma_5,$$

as explicitly shown by $\mathcal{H}'_0 = U_1^\dagger \mathcal{H}_0 U_1$ with U_1 defined by Eq. (18) in Ref. 2.

On the other hand, applying this transformation to $\mathcal{H}_m = (\mathbf{m} + \mathbf{n}\tau_z) \cdot \boldsymbol{\sigma}$ yields

$$\mathcal{H}'_m = (m_z + n_z \tau_z)\sigma_z + (\mathbf{n}_{||} + \mathbf{m}_{||}\tau_z) \cdot \boldsymbol{\sigma}.$$

Thus, a symmetric (antisymmetric) IP field in this paper corresponds to an antisymmetric (symmetric) one for the basis chosen in Ref. 2, while symmetric and antisymmetric OP fields are not switched by this transformation.

Appendix B: QSH phase

In the absence of $m_{||}$, the QSH state in thin-film TIs is simply defined by the opposite Chern numbers of two blocks classified by $p = \pm 1$, which correspond to the eigenvalues of $P = s_x \sigma_z$. Since $[P, H_0] = 0$, p is a conserved quantity.[12] However, it should be noted that the QSH state is not protected by the P-symmetry $PH_0 P^{-1} = H_0$. This is obvious from the phase transition into the QAH state due to a strong enough OP field.

As clearly shown in Fig. 1 (b), the bulk gap remains open in the QSH phase even for $m_{||} \neq 0$. The observation implies the nontrivial bulk topology persists, since the Chern numbers cannot change for two insulating states adiabatically connected without closing the gap. Analytically, this can be shown from the QSH criteria in the
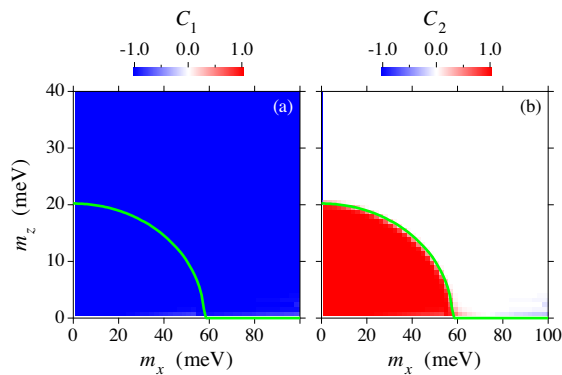
FIG. 6. (Color online) Numerical values of Chern numbers (a) C_1 and (b) C_2 for two valence bands ($E_1 \leq E_2$). Solid lines represent the phase boundaries where the gap closes. Parameters are identical to those of Fig. 1 (d), where $C = C_1 + C_2$ is shown.
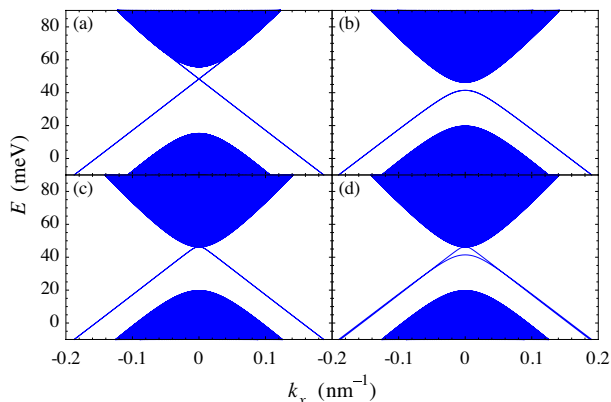

FIG. 7. (Color online) Energy dispersions computed in a strip geometry for (a) $m_x = 0$, (b) $m_x = m = 20$meV, (c) $m_x(y) = m\theta(\frac{L_y}{2} - w - |y|)$ and (d) $m_x(y) = m\theta(\frac{L_y}{2} - w - y)$. Parameters are $L_z = 3$nm, $L_y = 2000$nm, $m_y = m_z = 0$ and $w = 30$nm.

presence of a small $m_{||}$. Following the effective Hamil-

tonian, Eq. (7), the system belongs to the QSH phase as long as $\tilde{m}_z^2 < \tilde{\Delta}^2$ and $\tilde{\Delta}\tilde{B} > 0$. These criteria are entirely verified in the QSH phase shown in Fig. 1 (d). Based on Eqs. (3), (4) and (8), the opposite Chern numbers $C_1 = -C_2$ are numerically confirmed for two valence bands ($E_1 \leq E_2$) in the QSH phase, as shown in Fig. 6. Note that gap closing and reopening change the Chern number of the associated higher band.

However, the resulting QSH effect is rather complicated in a finite system with open boundaries. Figure 7 shows the energy spectra computed in a strip geometry of width L_y with a periodic boundary condition in the x direction and an open boundary condition in the y direction. In the calculation, parameters are chosen to be $L_z = 3$nm, $L_y = 2000$nm and $m_y = m_z = 0$. In the finite system, gapless helical edge modes exist for $m_x = 0$ [Fig. 7 (a)], while the edge modes are gapped by $m_x \equiv m = 20$meV [Fig. 7 (b)].

To examine the gap-opening mechanism, we have performed the additional calculations that assume a nonuniform field $m_x(y)$ partially vanishing in narrow regions adjacent to the boundaries at $y = \pm L_y/2$. The width of the field-free region is set at $w = 30$nm, which is slightly larger than the exponential decay length of edge-state wavefunction $\ell = 26$nm for $m_x = 0$.[12] Note that interedge interaction is negligible since $\ell << L_y$. Since $w << L_y$, the following two results are naturally expected. (i) The edge gap is sustained even for $m_x(y)$ if it reflects the bulk topology. (ii) The edge gap is suppressed for $m_x(y)$ if it arises from local intraedge interaction. The numerical results support the latter. The edge gap almost disappears for $m_x(y) = m\theta(\frac{L_y}{2} - w - |y|)$ [Fig. 7 (c)]. Moreover, for $m_x(y) = m\theta(\frac{L_y}{2} - w - y)$, the lower edge mode subjected to $m_x \neq 0$ is gapped while the upper edge mode little affected by m_x tends to be gapless [Fig. 7 (d)]. These results confirm that in the QSH phase, intraedge mixing due to $m_{||}$ creates the edge gap in a finite system, whereas the nontrivial bulk topology survives.

[1] H. Zhang, C.-X. Liu, X.-L. Qi, X. Dai, Z. Fang, and S.-C. Zhang, Nature Phys. **5**, 438 (2009).

[2] C.-X. Liu, X.-L. Qi, H. Zhang, X. Dai, Z. Fang, and S.-C. Zhang, Phys. Rev. B **82**, 045122 (2010).

[3] M. Z. Hasan and C. L. Kane, Rev. Mod. Phys. **82**, 3045 (2010).

[4] X.-L. Qi and S.-C. Zhang, Rev. Mod. Phys. **83**, 1057 (2011).

[5] J. Linder, T. Yokoyama, and A. Sudbø, Phys. Rev. B **80**, 205401 (2009).

[6] C.-X. Liu, H. Zhang, B. Yan, X.-L. Qi, T. Frauenheim, X. Dai, Z. Fang, and S.-C. Zhang, Phys. Rev. B **81**, 041307(R) (2010).

[7] H.-Z. Lu, W.-Y. Shan, W. Yao, Q. Niu, and S.-Q. Shen, Phys. Rev. B **81**, 115407 (2010).

[8] W.-Y. Shan, H.-Z. Lu, and S.-Q. Shen, New J. Phys. **12**, 043048 (2010).

[9] Y. Sakamoto, T. Hirahara, H. Miyazaki, S.-I. Kimura, and S. Hasegawa, Phys. Rev. B **81**, 165432 (2010).

[10] Y. Zhang, K. He, C.-Z. Chang, C.-L. Song, L.-L. Wang, X. Chen, J.-F. Jia, Z. Fang, X. Dai, W.-Y. Shan, S.-Q. Shen, Q. Niu, X.-L. Qi, S.-C. Zhang, X.-C. Ma, and Q.-K. Xue, Nature Phys. **6**, 584 (2010).

[11] H. Li, L. Sheng, D. N. Sheng, and D. Y. Xing, Phys. Rev. B **82**, 165104 (2010).

[12] K. Hattori, J. Phys. Soc. Jpn. **84**, 044701 (2015).

[13] C. L. Kane and E. J. Mele, Phys. Rev. Lett. **95**, 146802 (2005).

[14] B. A. Bernevig, T. L. Hughes, and S.-C. Zhang, Science **314**, 1757 (2006).

[15] M. König, S. Wiedmann, C. Brüne, A. Roth, H. Buhmann, L. W. Molenkamp, X.-L. Qi, and S.-C. Zhang, Science

318, 766 (2007).

[16] K. Hattori, J. Phys. Soc. Jpn. **80**, 124712 (2011).

[17] Y. S. Hor, P. Roushan, H. Beidenkopf, J. Seo, D. Qu, J. G. Checkelsky, L. A. Wray, D. Hsieh, Y. Xia, S.-Y. Xu, D. Qian, M. Z. Hasan, N. P. Ong, A. Yazdani, and R. J. Cava, Phys. Rev. B **81**, 195203 (2010).

[18] Y. L. Chen, J.-H. Chu, J. G. Analytis, Z. K. Liu, K. Igarashi, H.-H. Kuo, X. L. Qi, S. K. Mo, R. G. Moore, D. H. Lu, M. Hashimoto, T. Sasagawa, S. C. Zhang, I. R. Fisher, Z. Hussain, and Z. X. Shen, Science **329**, 659 (2010).

[19] S.-Y. Xu, M. Neupane, C. Liu, D. Zhang, A. Richardella, L. A. Wray, N. Alidoust, M. Leandersson, T. Balasubramanian, J. Sánchez-Barriga, O. Rader, G. Landolt, B. Slomski, J. H. Dil, J. Osterwalder, T.-R. Chang, H.-T. Jeng, H. Lin, A. Bansil, N. Samarth, and M. Z. Hasan, Nature Phys. **24**, 616 (2012).

[20] J. Zhang, C.-Z. Chang, P. Tang, Z. Zhang, X. Feng, K. Li, L.-L. Wang, X. Chen, C. Liu, W. Duan, K. He, Q.-K. Xue, X. Ma, and Y. Wang, Science **339**, 1582 (2013).

[21] C.-Z. Chang, P. Tang, Y.-L. Wang, X. Feng, K. Li, Z. Zhang, Y. Wang, L.-L. Wang, X. Chen, C. Liu, W. Duan, K. He, X.-C. Ma, and Q.-K. Xue, Phys. Rev. Lett. **112**, 056801 (2014).

[22] X.-L. Qi, Y.-S. Wu, and S.-C. Zhang, Phys. Rev. B. **74**, 085308 (2006).

[23] R. Yu, W. Zhang, H.-J. Zhang, S.-C. Zhang, X. Dai, and Z. Fang, Science **329**, 61 (2010).

[24] H.-Z. Lu, A. Zhao, and S.-Q. Shen, Phys. Rev. Lett. **111**, 146802 (2013).

[25] K. Hattori, J. Phys. Soc. Jpn. **83**, 034704 (2014).

[26] S.-F. Zhang, H. Jiang, X. C. Xie, and Q.-F. Sun, Phys. Rev. B **89**, 155419 (2014).

[27] C.-Z. Chang, J. Zhang, X. Feng, J. Shen, Z. Zhang, M. Guo, K. Li, Y. Ou, P. Wei, L.-L. Wang, Z.-Q. Ji, Y. Feng, S. Ji, X. Chen, J. Jia, X. Dai, Z. Fang, S.-C. Zhang, K. He, Y. Wang, L. Lu, X.-C. Ma, and Q.-K. Xue, Science **340**, 167 (2013).

[28] J. G. Checkelsky, R. Yoshimi, A. Tsukazaki, K. S. Takahashi, Y. Kozuka, J. Falson, M. Kawasaki, and Y. Tokura, Nature Phys. **10**, 731 (2014).

[29] X. Kou, S.-T. Guo, Y. Fan, L. Pan, M. Lang, Y. Jiang, Q. Shao, T. Nie, K. Murata, J. Tang, Y. Wang, L. He, T.-K. Lee, W.-L. Lee, and K. L. Wang, Phys. Rev. Lett. **113**, 137201 (2014).

[30] C.-Z. Chang, W. Zhao, D. Y. Kim, H. Zhang, B. A. Assaf, D. Heiman, S.-C. Zhang, C. Liu, M. H. W. Chan, and J. S. Moodera, Nature Mater. **14**, 473 (2015).

[31] C.-X. Liu, X.-L. Qi, X. Dai, Z. Fang, and S.-C. Zhang, Phys. Rev. Lett. **101**, 146802 (2008).

[32] C. Fang, M. J. Gilbert, and B. A. Bernevig, Phys. Rev. Lett. **112**, 046801 (2014).

[33] H. Zhang, Y. Xu, J. Wang, K. Chang, and S.-C. Zhang, Phys. Rev. Lett. **112**, 216803 (2014).

[34] Q.-Z. Wang, X. Liu, H.-J. Zhang, N. Samarth, S.-C. Zhang, and C.-X. Liu, Phys. Rev. Lett. **113**, 147201 (2014).

[35] S.-C. Wu, G. Shan, and B. Yan, Phys. Rev. Lett. **113**, 256401 (2014).

[36] H.-C. Hsu, X. Liu, and C.-X. Liu, Phys. Rev. B **88**, 085315 (2013).

[37] D. J. Thouless, M. Kohmoto, M. P. Nightingale, and M. den Nijs, Phys. Rev. Lett. **49**, 405 (1982).

[38] X. Liu, H.-C. Hsu, and C.-X. Liu, Phys. Rev. Lett. **111**, 086802 (2013).

[39] Y. Millev and J. Kirschner, Phys. Rev. B **54**, 4137 (1996).

[40] D. Stickler, R. Frömter, H. Stillrich, C. Menk, H. P. Oepen, C. Gutt, S. Streit-Nierobisch, L.-M. Stadler, and G. Grübel, Phys. Rev. B **84**, 104412 (2011).